
CHAPTER SEVEN: GLOSSARY

<i>Accountability</i>	The responsibility for the justification of expenditures, decisions, or the results of one's own efforts.
<i>Accuracy</i>	The extent to which an evaluation is truthful or valid in what it says about a program, project or material.
<i>Achievement</i>	A manifested performance determined by some type of assessment or testing.
<i>Adversarial/advocacy group</i>	A group of people who enter into cross-examination of counter plans, strategies, or outcomes.
<i>Affective</i>	Consists of emotions, feelings, and attitudes.
<i>Algorithm</i>	A step-by-step problem-solving procedure.
<i>Anonymity (provision for)</i>	Evaluator action to ensure that the identity of subjects cannot be ascertained during the course of a study, in study reports, or in any other way.
<i>Assessment</i>	Often used as a synonym for evaluation. The term is sometimes recommended for restriction to processes that are focused on quantitative and/or testing approaches.
<i>Attitude</i>	A person's mental set toward another person, thing, or state.
<i>Attrition</i>	Loss of subjects from the defined sample during the course of a longitudinal study.
<i>Audience(s)</i>	Consumers of the evaluation; those who will or should read or hear of the evaluation, either during or at the end of the evaluation process. Includes those persons who will be guided by the evaluation in making decisions and all others who have a stake in the evaluation (see stakeholders).
<i>Background</i>	The contextual information that describes the reasons for the project, its goals, objectives, and stakeholders' information needs.
<i>Baseline</i>	Facts about the condition or performance of subjects prior to treatment or intervention.

<i>Behavioral objectives</i>	Specifically stated terms of attainment to be checked by observation, or test/measurement.
<i>Bias</i>	A consistent alignment with one point of view.
<i>Case Study</i>	An intensive, detailed description and analysis of a single project, program, or instructional material in the context of its environment.
<i>Checklist approach</i>	Checklists are the principal instrument for practical evaluation; especially for investigating the thoroughness of implementation.
<i>Client</i>	The person or group or agency that commissioned the evaluation.
<i>Coding</i>	To translate a given set of data or items into machine-readable categories
<i>Cognitive</i>	The domain of knowledge—"knowledge-that" or "knowledge-how."
<i>Cohort</i>	A term used to designate one group among many in a study. For example, "the first cohort" may be the first group to have participated in a training program.
<i>Comparison group</i>	A group that provides a basis for contrast with (in experimentation) an experimental group (i.e., the group of people participating in the program or project being evaluated). The comparison group is not subjected to the treatment (independent variable), thus creating a means for comparison with the experimental group that does receive the treatment. Comparison groups should be "comparable" to the treatment group, but can be used when close matching is not possible (see also Control Group).
<i>Component</i>	A physically or temporally discrete part of a whole. It is any segment that can be combined with others to make a whole.
<i>Conceptual scheme</i>	A set of concepts that generate hypotheses and simplify description.
<i>Conclusions (of an evaluation)</i>	Final judgments and recommendations.
<i>Content analysis</i>	A process of systematically determining the characteristics of a body of material or practices.

<i>Control group</i>	A group that does not receive the treatment (service or product). The function of the control group is to determine the extent to which the same effect occurs without the treatment. The control group must be closely matched to the experimental group.
<i>Correlation</i>	A statistical measure of the degree of relationship between variables.
<i>Cost analysis</i>	The practical process of calculating the cost of something that is being evaluated. Cost analysis looks at: (1) costs to whom; (2) costs of what type; and (3) costs during what period.
<i>Cost-benefit analysis</i>	This process estimates the overall cost and benefit of each alternative product or program.
<i>Cost-effectiveness</i>	This analysis determines what a program or procedure costs against what it does (effectiveness). Is this product or program worth its costs?
<i>Criterion, criteria</i>	A criterion (variable) is whatever is used to measure as success, e.g., grade point average.
<i>Criterion-referenced test</i>	Tests whose scores are interpreted by referral to well defined domains of content or behaviors, rather than by referral to the performance of some comparable group of people.
<i>Cross-sectional study</i>	A cross-section is a random sample of a population, and a cross-sectional study examines this sample at one point in time. Successive cross-sectional studies can be used as a substitute for a longitudinal study. For example, examining today's first year students and today's graduating seniors may enable the evaluator to infer that the college experience has produced or can be expected to accompany the difference between them. The cross sectional study substitutes today's seniors for a population that cannot be studied until 4 years later.
<i>Delivery system</i>	The link between the product or service and the immediate consumer (the recipient population).
<i>Dependent variable</i>	One that represents the outcome—the contrast is with independent variables some of which can be manipulated.
<i>Descriptive statistics</i>	Those that involve summarizing, tabulating, organizing, and graphing data for the purpose of describing objects or individuals that have been measured or observed.

<i>Design</i>	The process of stipulating the investigatory procedures to be followed in doing a certain evaluation.
<i>Dissemination</i>	The process of communicating information to specific audiences for the purpose of extending knowledge and, in some cases, with a view to modifying policies and practices.
<i>Effectiveness</i>	Refers to the conclusion of a Goal Achievement Evaluation. “Success” is its rough equivalent.
<i>Executive report</i>	An abbreviated report that has been tailored specifically to address the concerns and questions of a person whose function is to administer an educational program or project.
<i>Executive summary</i>	A nontechnical summary statement designed to provide a quick overview of the full-length report on which it is based.
<i>Experimental design</i>	The plan of an experiment, including selection of subjects who receive treatment and control group (if applicable), procedures, and statistical analyses to be performed.
<i>Experimental group</i>	The group that is receiving the treatment.
<i>External evaluation</i>	Evaluation conducted by an evaluator from outside the organization within which the object of the study is housed.
<i>Extrapolate</i>	To infer an unknown from something that is known. (Statistical definition—to estimate the value of a variable outside its observed range.)
<i>False positive</i>	When an event is predicted and it does not occur (Type I error).
<i>False negative</i>	When an event is not predicted and it occurs (Type II error).
<i>Feasibility</i>	The extent to which an evaluation is appropriate for implementation in practical settings.
<i>Field test</i>	The study of a program, project, or instructional material in settings like those where it is to be used. Field tests may range from preliminary primitive investigations to full-scale summative studies.

<i>Flow chart</i>	A graphic representation of a set of decisions that is set up to guide the management of projects, including evaluation projects.
<i>Focus group</i>	A group selected for its relevance to an evaluation that is engaged by a trained facilitator in a series of discussions designed for sharing insights, ideas, and observations on a topic of concern.
<i>Formative evaluation</i>	Evaluation designed and used to improve an intervention, especially when it is still being developed.
<i>Gain scores</i>	The difference between a student's performance on a test and his or her performance on a previous administration of the same or parallel test.
<i>Generalizability</i>	The extent to which information about a program, project, or instructional material collected in one setting can be used to reach a valid judgment about how it will perform in other settings.
<i>Goal-free evaluation</i>	Evaluation of outcomes in which the evaluator functions without knowledge of the purposes or goals.
<i>Hawthorne effect</i>	The tendency of a person or group being investigated to perform better (or worse) than they would in the absence of the investigation, thus making it difficult to identify treatment effects.
<i>Hypothesis testing</i>	The standard model of the classical approach to scientific research in which a hypothesis is formulated before the experiment to test its truth. The results are stated in probability terms that the results were due solely to chance. The significance level of one chance in 20 (.05) or one chance in 100 (.01) is a high degree of improbability.
<i>Impact evaluation</i>	An evaluation focused on outcomes or pay-off.
<i>Implementation evaluation</i>	Assessing program delivery (a subset of Formative Evaluation).
<i>Indicator</i>	A factor, variable, or observation that is empirically connected with the criterion variable, a correlate. For example, judgment by students that a course has been valuable to them for pre-professional training is an indicator of that value.
<i>Inferential statistics</i>	These statistics are inferred from characteristics of samples to characteristics of the population from which the sample comes.

<i>Informed consent</i>	Agreement by the participants in an evaluation of the use of their names and/or confidential information supplied by them in specified ways, for stated purposes, and in light of possible consequences prior to the collection and/or release of this information in evaluation reports.
<i>Instrument</i>	An assessment device (test, questionnaire, protocol, etc.) adopted, adapted, or constructed for the purpose of the evaluation.
<i>Interaction</i>	Two factors or variables interact if the effect of one, on the phenomenon being studied, depends upon the magnitude of the other. For example, mathematics education interacts with age, being more or less effective depending upon the age of the child.
<i>Internal evaluator</i>	Internal evaluations are those done by project staff, even if they are special evaluation staff, that is, external to the production/writing/ teaching/service part of the project.
<i>Level of significance</i>	The probability that the observed difference occurred by chance.
<i>Longitudinal study</i>	An investigation or study in which a particular individual or group of individuals is followed over a substantial period of time to discover changes due to the influence of the treatment, or maturation, or environment.
<i>Mastery level</i>	The level of performance needed on a criterion. The mastery level is often arbitrary.
<i>Matching</i>	An experimental procedure in which the subjects are so divided, by means other than lottery, that the groups are regarded for the purposes at hand to be of equal merit or ability. (Often matched groups are created by ensuring that they are the same or nearly so on such variables as sex, age, grade point averages, and past test scores.)
<i>Matrix</i>	An arrangement of rows and columns used to display components of evaluation design.
<i>Mean</i>	Also called "average" or arithmetic average. For a collection of raw test scores, the mean score is obtained by adding all scores and dividing by the number of people taking the test.

Measurement	Determination of the magnitude of a quantity.
Median	The point in a distribution which divides the group into two, as nearly as possible. For example, in a score distribution, half the scores fall above the median and half fall below.
Meta-analysis	The name for a particular approach to synthesizing quantitative studies on a common topic, involving the calibration of a specific parameter for each ("effect size").
Metric data	Data which includes a unit of measurement (i.e., dollars, inches).
Mode	The value which occurs more often than any other. If all scores (in a score distribution) occur with the same frequency, there is no mode. If the two highest score values occur with the same frequency, there are two modes.
Needs assessment	Using a diagnostic definition, need is anything essential for a satisfactory mode of existence or level of performance. The essential point of a needs assessment for evaluation is the identification of performance needs.
Nominal data	Data which consist of categories only without order to these categories (i.e., region of the country, courses offered by an instructional program).
"No significant difference"	A decision that an observed difference between two statistics occurred by chance.
Nonreactive measures	Assessments done without the awareness of those being assessed.
Norm	A single value, or a distribution of values, constituting the typical performance of a given group.
Norm-referenced tests	Tests that measure the <i>relative</i> performance of the individual or group by comparison with the performance of other individuals or groups taking the same test.
Objective	A specific description of an intended outcome.
Observation	The process of direct sensory inspection involving trained observers.
Operational definition	A definition of a term or object achieved by stating the operations or procedures employed to distinguish it from others.

<i>Ordered data</i>	Non-numeric data in ordered categories (for example, students performance being categorized as excellent, good, adequate, and poor).
<i>Outcome</i>	Post-treatment or post-intervention effects.
<i>Paradigm</i>	A general conception of or model for a discipline or subdiscipline which may be very influential in shaping its development. (For example, “The classical social science paradigm in evaluation.”)
<i>Peer review</i>	Evaluation done by a panel of judges with qualifications approximating those of the author or candidate.
<i>Performance-based</i>	The use of global ratings of behavior assessment which is a movement away from paper-and-pencil testing. This assessment is costly and there may be a loss of validity and reliability.
<i>Pilot test</i>	A brief and simplified preliminary study designed to try out methods to learn whether a proposed project or program seems likely to yield valuable results.
<i>Planning evaluation</i>	Evaluation planning is necessary before a program begins, both to get baseline data, and to evaluate the program plan, at least for evaluability. Planning avoids designing a program that is unevaluable.
<i>Population</i>	All persons in a particular group.
<i>Post test</i>	A test to determine performance after the administration of a program, project, or instructional material.
<i>Pretest</i>	A test to determine performance prior to the administration of a program, project, or instructional material. Pretests serve two purposes: diagnostic and baseline. Also the use of an instrument (questionnaire, test, observation schedule) with a small group to detect need for revisions.
<i>Process evaluation</i>	Refers to the evaluation of the treatment or intervention. It focuses entirely on the variables between input and output.
<i>Product</i>	A pedagogical process or material coming from research and development.
<i>Program</i>	The general effort that marshals staff and projects toward defined and funded goals.
<i>Progress evaluation</i>	A subset of Formative Evaluation.

Prompt	Reminders used by interviewers to obtain complete answers.
Qualitative evaluation	The part of the evaluation that is primarily descriptive and interpretative, and may or may not lend itself to quantitative treatment.
Quantitative evaluation	An approach involving the use of numerical measurement and data analysis based on statistical methods.
Quasi-experimental	When a random allocation of subjects to experimental and control groups can not be done, a quasi-experimental design can seek to simulate a true experimental design, by identifying a group that closely matches the experimental group.
Random	Affected by chance.
Random sampling	Drawing a number of items of any sort from a larger group or population so that every individual item has a specified probability of being chosen.
Recommendations	Suggestions for specific appropriate actions based upon analytic approaches to the program components.
Reliability	Statistical reliability is the consistency of the readings from a scientific instrument or human judge.
Remediation	The process of improvement or a recommendation for a course of action or treatment that will result in improvement.
Replication	Repeating an intervention or evaluation with all essentials unchanged. Replications are often difficult to evaluate because of changes in design or execution.
Research	The general field of disciplined investigation.
Response bias	Error due to incorrect answers.
Sample	A part of a population.
Sample bias	Error due to non-response or incomplete response from selected sample subjects.
Sampling error	Error due to using a sample instead of entire population from which sample is drawn.
Secondary data analysis	A reanalysis of data using the same or other appropriate procedures to verify the accuracy of the results of the initial analysis or for answering different questions.

<i>Self-administered instrument</i>	A questionnaire or report completed by a study participant without the assistance of an interviewer.
<i>Self-report instrument</i>	A device in which persons make and report judgments about the functioning of their project, program, or instructional material.
<i>Significance</i>	Overall significance represents the total <i>synthesis</i> of all you have learned about the merit or worth of the program or project. This is different from statistical significance which may be testing one of several conditions of a program or project.
<i>Stakeholder</i>	A program's stakeholder is one who has credibility, power, or other capital invested in the project, and thus can be held to be to some degree at risk with it.
<i>Standard deviation</i>	A measure of the spread of a variable, based on deviation from the mean value for metric data.
<i>Standardized tests</i>	Tests that have standardized instructions for administration, use, scoring, and interpretation with standard printed forms and content. They are usually norm-referenced tests but can also be criterion-referenced.
<i>Statistic</i>	A summary number that is typically used to describe a characteristic of a sample.
<i>Strategy</i>	A systematic plan of action to reach predefined goals.
<i>Summary</i>	A short restatement of the main points of a report.
<i>Summative evaluation</i>	Evaluation designed to present conclusions about the merit or worth of an intervention and recommendations about whether it should be retained, altered, or eliminated.
<i>Time series study</i>	A study in which periodic measurements are obtained prior to, during, and following the introduction of an intervention or treatment in order to reach conclusions about the effect of the intervention.
<i>Treatment</i>	Whatever is being investigated; in particular, whatever is being applied or supplied to, or done by, the experimental groups that is intended to distinguish them from the comparison groups.
<i>Triangulation</i>	In an evaluation, it is an attempt to get a fix on a phenomenon or measurement by approaching it via several independent routes. It can be more than three routes. This effort provides redundant measurement.

<i>Unanticipated outcomes</i>	A result of a program or interview that was unexpected. Often used as a synonym for side-effects, but only a loose equivalent.
<i>Utility</i>	The extent to which an evaluation produces and disseminates reports that inform relevant audiences and have beneficial impact on their work.
<i>Utilization (of evaluations)</i>	Use and impact are terms used as substitutes for utilization. Sometimes seen as the equivalent of implementation, but this applies only to evaluations which contain recommendations.
<i>Validity</i>	The soundness of the use and interpretation of a measure.

SOURCES

Jaeger, R. M. (1990). *Statistics A Spectator Sport*. Newbury Park, CA: Sage.

Joint Committee on Standards for Educational Evaluations (1981). *Standards for Evaluation of Educational Programs, Projects and Materials*. New York, NY: McGraw Hill.

Rogers, E. M. (1983). *Diffusion of Innovations*. New York, NY: Free Press.

Scriven, M. (1991). *Evaluation Thesaurus*. Fourth Edition. Newbury Park, CA: Sage.

Authors of Chapters 1-4.

